NOVEMBER/DECEMBER 2025

23PCS21 — DATA MINING AND WAREHOUSING

Time : Three hours Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL the questions.

1. Differentiate — Classification and Prediction.
2. Mention any two Data mining issues.
3. What is Propagation in Neural network?
4. What are Rule-based algorithms?
5. Define the term — Clustering.
6. What are Association rules?
7. Define the term — OLTP Systems.
8. What is Snow Flake Schema?
9. Define the term — Data Warehouse.
10. What is the purpose of Data warehouse?

Shanmuga Industries Arts & Science College T.V.Malai

SECTION B — (5 × 5 = 25 marks)

Answer ALL the questions.

11. (a) Discuss on — Data mining versus Knowledge discovery in Databases.

Or

(b) Explain the concept of Decision trees.

12. (a) Describe about Distance-based algorithms in Classification.

Or

(b) Discuss on ID3 and C4.5 Decision tree-based algorithms in Classification.

13. (a) Explain the basics of Similarity, Distance measures, and Outliers.

Or

(b) Describe about K-Means Clustering algorithms.

14. (a) Discuss on Star Schema for Multidimensional View in Data modeling.

Or

(b) Write short notes on OLAP Tools and the Internet.

Shanmuga Industries Arts & Science College T.V.Malai

15. (a) Explain the Design considerations, Data content and Meta data in Data warehouse design.

Or

(b) Describe about the Tools for Data warehousing and the performance considerations.

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Describe about the Statistical perspective on Data mining.

17. Explain the Statistical-based Algorithms in Classification in detail.

18. Discuss on the working of Apriori algorithm in Association rule mining.

19. Describe the fundamentals of Data Marts in detail.

20. Explain the National Data Warehouses, and other Areas for Data warehousing and Data mining.

———————